ABSTRACT

Automatic interpretation processing operations, such as character recognition, require a binary image of information-bearing image elements and a background. Digital image data produced by scanning a colour-containing document frequently contain many different colours so that a division into information-bearing image elements and background is not evident. The invention divides connected components (contiguous pixels with the same colour) in the digital image into background and other connected components, whereafter the other connected components are allocated, in accordance with a predetermined criterion, either to the background or to a foreground. The foreground connected components are combined into information elements suitable for the automatic interpretation processing. The division into information-bearing image elements and background can be preceded by a colour quantisation processing of the digital image data.